REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, seraching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washing Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503

1. AGENCY USE ONLY *(Leave blank)*

2. REPORT DATE 8/27/98

3. REPORT TYPE AND DATES COVERED
Final Report (3/1/95 through 2/28/98)

4. TITLE AND SUBTITLE
Final Report: Hybrid Computational Models for Skill Acquisition

5. FUNDING NUMBERS
N00014-95-1-0557

6. AUTHOR(S)
Thomas G. Dietterich
Prasad Tadepalli

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Department of Computer Science
Oregon State University
Corvallis, Oregon 97331

8. PERFORMING ORGANIZATION REPORT NUMBER
None

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Office of Naval Research
800 North Quincy Street, Ballston Tower One
Arlington, VA 22217-5660

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

Approved for public release
Distribution Unlimited

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Distribution Unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

The goal of this research was to develop a hybrid real-time problem-solving architecture that couples symbolic planning methods with connectionist reinforcement learning methods. The advantage of this hybrid architecture is that it can immediately achieve reasonable performance, because the symbolic planning system can quickly develop an acceptable control policy, but it can also gradually achieve optimal real-time performance, because the reinforcement learning system will eventually converge on a near-optimal policy. Many DoD problems would benefit from the ability to perform near-optimal real-time control of complex systems.

DTIC QUALITY INSPECTED 1

14. SUBJECT TERMS
real-time problem solving, machine learning, reinforcement learning, planning

15. NUMBER OF PAGES
5

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
unclassified	unclassified	unclassified	

Final Report:
Hybrid Computational Methods for Skill Acquisition
ONR Grant Number N00014-95-1-0557

Thomas G. Dietterich
Prasad Tadepalli

Department of Computer Science
303 Dearborn Hall
Oregon State University
Corvallis, OR 97331

August 27, 1998.

19980904 083

1 Administrative Information

Title:	Hybrid Computational Methods for Skill Acquisition
Grant Number:	N00014-95-1-0557
Grantee:	Oregon State University, Corvallis, OR 97331
Start Date:	March 1, 1995
End Date:	February 28, 1998
Principal Investigators:	Thomas G. Dietterich Department of Computer Science 303 Dearborn Hall Oregon State University Corvallis, OR 97331 503-737-5559 (voice) 503-737-3014 (fax) `tgd@cs.orst.edu` (email) Prasad Tadepalli Department of Computer Science 303 Dearborn Hall Oregon State University Corvallis, OR 97331 503-737-5552 (voice) 503-737-3014 (fax) `tadepall@cs.orst.edu` (email)

DEPARTMENT OF
COMPUTER SCIENCE

COPY

FYI

August 28, 1998

Dr. Michael Shneier
Office of Naval Research
800 North Quincy Street
Ballston Tower One
Arlington, VA 22217-5660

OREGON STATE UNIVERSITY

Dearborn Hall 303
Corvallis, Oregon
97331-3202

Dear Michael Shneier:

Please find enclosed two copies of our final report for N00014-95-1-0557. I have also sent one copy to the Defense Technical Information Center. If there are other people who should receive copies, please let me know. This report contains the same information that we sent in earlier via electronic mail.

All publications supported by this grant are available for downloading from our web pages.

Thank-you very much for your support of this research. It has been a very successful project.

Sincerely,

Thomas G. Dietterich
Professor

Telephone
541 · 737 · 3273
Fax
541 · 737 · 3014

2 Project Summary

The goal of this research was to develop a hybrid real-time problem-solving architecture that couples symbolic planning methods with connectionist reinforcement learning methods. The advantage of this hybrid architecture is that it can immediately achieve reasonable performance, because the symbolic planning system can quickly develop an acceptable control policy, but it can also gradually achieve optimal real-time performance, because the reinforcement learning system will eventually converge on a near-optimal policy. Many DoD problems would benefit from the ability to perform near-optimal real-time control of complex systems.

3 Accomplishments

- Developed the ALERT hybrid architecture which combines symbolic (DRULE) planner with hierarchical reinforcement learning
- Showed experimentally that the DRULE planner could achieve human-level performance on the Kanfer-Ackerman air traffic control (ATC) task.
- Developed two learning algorithms for DRULES: one based on random examples and queries, and the other based on exercises.
- Showed experimentally that both learning algorithms could achieve intermediate performance on the ATC task.
- Proved that both learning algorithms are correct and computationally feasible. This involved proving a new result on learning of Horn clause logic programs.
- Developed a new, hierarchical method for reinforcement learning, the MAXQ method.
- Proved that MAXQ can represent any hierarchical policy.
- Developed the MAXQ-Q learning algorithm for hierarchical reinforcement learning.
- Proved that MAXQ-Q converges to a recursively optimal policy asymptotically.
- Demonstrated experimentally that MAXQ-Q attains optimal performance on a simplified task that shares many properties with the ATC task.

4 Transitions

We are currently working with i2 Technologies (Dallas, Texas) to apply our reinforcement learning methods to supply chain scheduling and optimization.

5 ONR-Funded Publications for entire grant period

Dietterich, T. G. (1998). The MAXQ Method for Hierarchical Reinforcement Learning. Proceedings of the Fifteenth International Conference on Machine Learning. Madison, Wisconsin. San Francisco: Morgan Kaufmann.

Dietterich, T. G. (1998). Hierarchical reinforcement learning with the MAXQ value function decomposition. Technical Report. Department of Computer Science. Oregon State University.

Reddy, C., Tadepalli, P. (1998). Learning First-Order Acyclic Horn Programs from Entailment. Proceedings of the Fifteenth International Conference on Machine Learning, Madison, Wisconsin. San Francisco: Morgan Kaufmann.

Reddy, C., Tadepalli, P. (In press). Learning Horn Definitions: Theory and an Application to Planning. New Generation Computing. 16(4).

Tadepalli, P. and Ok, D. Model-based Average Reward Reinforcement Learning. Artificial Intelligence, 100, 177-224, 1998.

Zhang, W., Dietterich, T. G. (1998). Solving Combinatorial Optimization Tasks by Reinforcement Learning: A General Methodology Applied to Resource-Constrained Scheduling. Technical Report. Department of Computer Science, Oregon State University.

Dietterich, T. G. (in press). Statistical Tests for Comparing Supervised Classification Learning Algorithms. Neural Computation.

Dietterich, T. G., Flann, N. S., (1997). Explanation-based Learning and Reinforcement Learning: A Unified View. Machine Learning, 28(2), 169–210.

Reddy, C. and Tadepalli, P. (1997) Learning Horn Definitions using Equivalence and Membership Queries. International Conference on Inductive Logic Programming.

Reddy, C. and Tadepalli, P. (1997). Learning goal-decomposition rules using exercises. In Proceedings of the Fourteenth International Conference on Machine Learning. 278-286. San Francisco: Morgan Kaufmann.

Reddy, C., Tadepalli, P. (1997) Inductive Logic Programming for Speedup Learning. Proceedings of the International Joint Conference on Artificial Intelligence Workshop on Frontiers of Inductive Logic Programming. Nagoya, Japan.

Reddy, C., Tadepalli, P. (1997) Learning Horn Programs using Equivalence and Membership Queries. Proceedings of the International Joint Conference on

Artificial Intelligence Workshop on Frontiers of Inductive Logic Programming. Nagoya, Japan.

Tadepalli, P., Dietterich, T. G. (1997). Hierarchical Explanation-Based Reinforcement Learning. In Proceedings of the Fourteenth International Conference on Machine Learning. San Francisco: Morgan Kaufmann. 358–366.

Zhang, W., Dietterich, T. G., (1996). High-Performance Job-Shop Scheduling With A Time-Delay $TD(\lambda)$ Network. In Advances in Neural Information Processing Systems, 8, 1024–1030.

Zhang, W., Dietterich, T. G., (1995). A Reinforcement Learning Approach to Job-shop Scheduling. In 1995 International Joint Conference on Artificial Intelligence (pp. 1114–1120) Montreal, Canada.

Dietterich, T. G., Flann, N. S., (1995). Explanation-based Learning and Reinforcement Learning: A Unified View. In Proceedings of the 12th International Conference on Machine Learning (pp. 176–184) Tahoe City, CA. San Francisco: Morgan Kaufmann.

6 Online Information Available

Postscript files for all papers are available via WWW from the following URL's:

```
http://www.cs.orst.edu/~tgd/
http://www.cs.orst.edu/~tadepall/
```